Analyse 3 - Dérivabilité

1 Dérivée en un point

Dans toute cette partie, f est une fonction définie au voisinage V de $a \in \mathbb{R}$ ($V = [a-\eta, a+\eta]$, avec $\eta > 0$) et à valeurs dans $\mathbb{R}$.
On note V_0 le voisinage de 0 correspondant : $V_0 = [-\eta, \eta]$.

Définition (Première définition de la dérivabilité)**.**
On dit que f est dérivable en a lorsque le taux d'accroissement $\frac{f(t)-f(a)}{t-a}$, défini pour $t \in V \setminus \{a\}$, admet une limite finie $d \in \mathbb{R}$ lorsque $t \to a$.
Lorsque cette limite d existe, on la note $f'(a)$.

Remarque. *Par le changement de variable $t = a + x$, la dérivabilité de f en a se ramène à l'étude de la limite lorsque $x \to 0$ de $\frac{f(a+x)-f(a)}{x}$.*

Exemple : étudions la dérivabilité de la fonction $f \ : \ t \mapsto \frac{1}{t}$ en un point $a \in \mathbb{R}^*$. Pour tout x appartenant à un voisinage V_0 de 0, on a :

$$f(a+x) - f(a) = \frac{1}{a+x} - \frac{1}{a} = \frac{a-(a+x)}{a(a+x)} = \frac{-x}{(a+x)a}$$

Donc pour $x \neq 0$:

$$\frac{f(a+x)-f(a)}{x} = \frac{-1}{a(a+x)} \underset{x\to 0}{\longrightarrow} \frac{-1}{a^2}$$

Cela prouve que f est dérivable en a, et que $\boxed{f'(a) = \frac{-1}{a^2}}$.

Remarque fondamentale

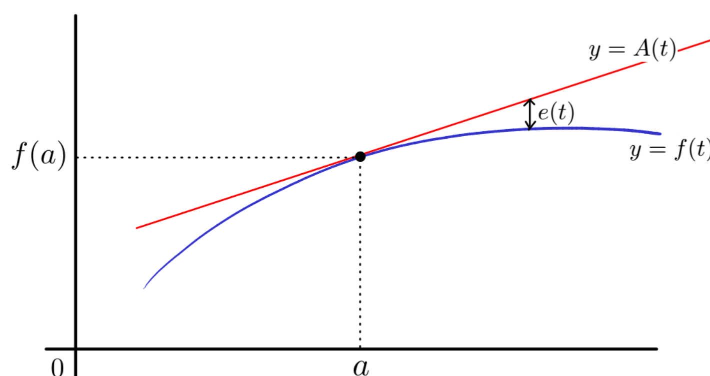
Dire que $\frac{f(t)-f(a)}{t-a}$ tend vers un réel d lorsque $t \to a$ est équivalent à dire que pour tout $t \in V$:

$$\frac{f(t)-f(a)}{t-a} = d + \varepsilon_a(t) \quad \text{où } \varepsilon_a \text{ est une fonction telle que } \varepsilon_a(t) \underset{t\to a}{\longrightarrow} 0.$$

Or, on a les équivalences suivantes :

$$\begin{aligned} \frac{f(t)-f(a)}{t-a} = d + \varepsilon_a(t) &\iff f(t) - f(a) = d(t-a) + (t-a)\varepsilon_a(t) \\ &\iff f(t) = \underbrace{f(a) + d\,(t-a)}_{A(t)} + \underbrace{(t-a)\varepsilon_a(t)}_{e(t)} \end{aligned}$$

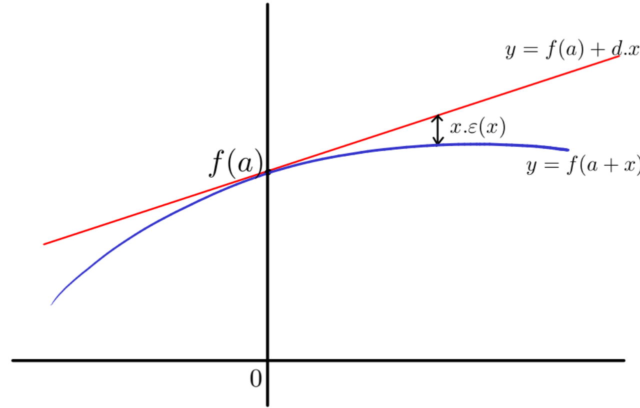
Dans cette dernière égalité :
- la quantité $A(t)$ est l'expression d'une fonction affine, et plus précisément de l'unique fonction affine de pente d et valant $f(a)$ en a ;
- la quantité $e(t)$ est une quantité qui tend vers 0 "plus vite que $(t-a)$" lorsque $t \to a$, au sens où elle est le produit de $(t-a)$ (qui tend vers 0) et de $\varepsilon_a(t)$ qui tend aussi vers 0.



Moralement, ces équivalences montrent qu'une fonction est dérivable en a si et seulement si elle admet une approximation par une fonction affine. On parle alors de "développement limité à l'ordre 1". Géométriquement, la fonction affine A définit la droite tangente à la courbe de f en a.

Avec le changement de variable $t = a + x$, c'est aussi équivalent à dire que pour $x \in V_0$, on a :

$$f(a+x) = f(a) + d.x + x\varepsilon(x) \quad \text{où } \varepsilon \text{ est une fonction telle que } \varepsilon(x) \underset{x\to 0}{\longrightarrow} 0$$



Définition (Seconde définition de la dérivabilité).
On dit que f est dérivable en a s'il existe $d \in \mathbb{R}$ et une fonction ε telle que :

$$\forall x \in V_0\,,\; f(a+x) = f(a) + d.x + x\varepsilon(x) \qquad \textit{et} \qquad \varepsilon(x) \underset{x\to 0}{\longrightarrow} 0$$

Le réel d est alors unique, et est noté $f'(a)$.

L'équivalence entre les deux définitions est montrée dans la remarque fondamentale.

Théorème. *Si f est dérivable en a, alors f est continue en a.*

Démonstration. Supposons que f est dérivable en a. Alors d'après la seconde définition de la dérivabilité, on a :

$$\forall x \in V_0\,,\; f(a+x) = f(a) + f'(a).x + x\varepsilon(x) \qquad \text{où } \varepsilon \text{ est une fonction telle que } \varepsilon(x) \underset{x\to 0}{\longrightarrow} 0$$

Lorsque $x \to 0$, on a $\begin{cases} f'(a).x \to 0 \\ x\varepsilon(x) \to 0 \end{cases}$ et donc $f(a+x) \underset{x\to 0}{\longrightarrow} f(a)$.
Cela signifie que $f(t) \underset{t\to a}{\longrightarrow} f(a)$, c'est-à-dire que f est continue en a. □

Proposition (Dérivation d'une composée). *Si u est une fonction dérivable en a et f une fonction dérivable en $u(a)$, alors $f \circ u$ est dérivable en a, et :*

$$(f \circ u)'(a) = f'(u(a)) \times u'(a)$$

Démonstration. D'après la seconde définition de la dérivabilité, il existe deux fonctions ε et E qui tendent vers 0 en 0, et telles que pour tous x, X appartenant à un voisinage de 0 :

$$\begin{cases} u(a+x) = u(a) + x\big(u'(a) + \varepsilon(x)\big) & (1) \\ f(u(a)+X) = f(u(a)) + X\big(f'(u(a)) + E(X)\big) & (2) \end{cases}$$

On applique l'égalité (2) à $X = x\big(u'(a) + \varepsilon(x)\big)$, qui est une quantité qui tend vers 0 lorsque $x \to 0$, de façon à ce que la quantité $f(u(a) + X)$ soit égale à $f(u(a + x))$. On obtient alors :

$$\begin{aligned} f(u(a+x)) &= f(u(a)) + x\big(u'(a) + \varepsilon(x)\big)\left(f'(u(a)) + E\left(x\big(u'(a)+\varepsilon(x)\big)\right)\right) \\ &= \underbrace{f(u(a))}_{f\circ u(a)} + x\underbrace{\big(f'(u(a)) \times u'(a)\big)}_{\text{formule annoncée}} + x\underbrace{(\ldots\ldots\ldots\ldots)}_{\text{termes qui } \underset{x\to 0}{\longrightarrow} 0} \end{aligned}$$

D'après la seconde définition de la dérivabilité, ce travail montre que $f \circ u$ est dérivable en a, et que $(f \circ u)'(a) = f'(u(a)) \times u'(a)$. □

Application : si u est une fonction dérivable en $a \in \mathbb{R}$, et telle que $u(a) \neq 0$, que peut-on dire de la dérivabilité de $\frac{1}{u}$ en a ?

La fonction $\frac{1}{u}$ peut être interprétée comme étant la fonction $f \circ u$, où $f : t \mapsto \frac{1}{t}$. Comme $u(a) \neq 0$, la fonction $f \circ u$ est bien définie en a. De plus, sait que :

- u est dérivable en a (par hypothèse) ;
- f est dérivable en $u(a)$ (vu précédemment dans ce cours) et $f'(u(a)) = \frac{-1}{u(a)^2}$.

D'après le théorème précédent, $f \circ u$ est donc dérivable en a, comme composée de fonctions dérivables, et :

$$\big(f(u)\big)'(a) = f'(u(a)) \times u'(a) = \frac{-u'(a)}{u(a)^2}$$

Bien sûr, cela s'applique *en tout point a* où u est dérivable (et tel que $u(a) \neq 0$), de sorte qu'on a bien démontré ici que sur tout intervalle sur lequel u est dérivable (et ne s'annule pas), la fonction $\frac{1}{u}$ est elle-même dérivable, et :

$$\left(\frac{1}{u}\right)' = \frac{-u'}{u^2}$$

Propositions (Opérations algébriques)**.** *Soit u et v deux fonctions définies sur V_a un voisinage de a, et dérivables en a. Alors :*

1. *$u + v$ est dérivable en a, et $(u + v)'(a) = u'(a) + v'(a)$;*
2. *$\forall \lambda \in \mathbb{R}$, λu est dérivable en a et $(\lambda u)'(a) = \lambda u'(a)$;*
3. *$u \times v$ est dérivable en a, et $(uv)'(a) = u'(a)v(a) + u(a)v'(a)$;*
4. *si $v(a) \neq 0$, $\frac{u}{v}$ est dérivable en a, et $\left(\frac{u}{v}\right)'(a) = \frac{u'(a)v(a) - u(a)v'(a)}{v(a)^2}$.*

Démonstration. On part de la dérivabilité de u et v en a : il existe deux fonctions ε_1 et ε_2 qui tendent vers 0 en 0 et telles que pour $x \in V_0$:

$$\begin{cases} u(a+x) = u(a) + u'(a).x + x.\varepsilon_1(x) \\ v(a+x) = v(a) + v'(a).x + x.\varepsilon_2(x) \end{cases}$$

De là, on a :

1. pour la somme :

$$\begin{aligned} (u+v)(a+x) &= u(a+x) + v(a+x) \\ &= \underbrace{\big(u(a) + v(a)\big)}_{(u+v)(a)} + \underbrace{\big(u'(a) + v'(a)\big)}_{\text{comme annoncé}}.x + x.\underbrace{\big(\varepsilon_1(x) + \varepsilon_2(x)\big)}_{\text{tend vers } 0} \end{aligned}$$

   D'après la seconde définition de la dérivabilité, cela démontre que $(u + v)$ est dérivable en a, et que $(u + v)'(a) = u'(a) + v'(a)$.

2. pour λu :

$$\begin{aligned} (\lambda u)(a+x) &= \lambda.u(a+x) \\ &= \lambda\big(u(a) + u'(a).x + x.\varepsilon_1(x)\big) \\ &= \underbrace{\big(\lambda.u(a)\big)}_{(\lambda u)(a)} + \underbrace{\big(\lambda.u'(a)\big)}_{\text{comme annoncé}}.x + x.\underbrace{\big(\lambda.\varepsilon_1(x)\big)}_{\text{tend vers } 0} \end{aligned}$$

   D'après la seconde définition de la dérivabilité, cela démontre que $(\lambda.u)$ est dérivable en a, et que $(\lambda.u)'(a) = \lambda.u'(a)$.

3. pour le produit :

$$\begin{aligned}(u\times v)(a+x) &= u(a+x)\times v(a+x)\\ &= \big(u(a)+u'(a).x+x.\varepsilon_1(x)\big)\times\big(v(a)+v'(a).x+x.\varepsilon_2(x)\big)\\ &= \underbrace{\big(u(a)\times v(a)\big)}_{(u\times v)(a)}+\underbrace{\big(u'(a)v(a)+u(a)v'(a)\big)}_{\text{comme annoncé}}.x+x.\underbrace{\big(\ldots\ldots\ldots\ldots\ldots\big)}_{\text{6 termes qui }\underset{x\to 0}{\longrightarrow}0}\end{aligned}$$

NB : la dernière ligne a été obtenue en distribuant le produit ($3\times 3=9$ termes).
D'après la seconde définition de la dérivabilité, cela démontre que $(u\times v)$ est dérivable en a, et que $(u\times v)'(a)=u'(a)v(a)+u(a)v'(a)$.

4. pour le quotient, on peut remarquer que $\frac{u}{v}=u\times\frac{1}{v}$. Or, le cas de $\frac{1}{v}$ a été éclairci par l'application faite avant ce théorème. La fonction $\frac{1}{v}$ est dérivable en a, et $(\frac{1}{v})'(a)=\frac{-v'(a)}{v^2(a)}$. Donc par produit, la fonction $u\times\frac{1}{v}$ est dérivable en a, et :

$$\begin{aligned}\left(u\times\frac{1}{v}\right)'(a) &= u'(a)\cdot\left(\frac{1}{v}\right)(a)+u(a)\cdot\left(\frac{1}{v}\right)'(a) && \text{(formule de dérivation du produit)}\\ &= \frac{u'(a)}{v(a)}+u(a)\cdot\left(\frac{-v'(a)}{v^2(a)}\right)\\ &= \frac{u'(a)v(a)-u(a)v'(a)}{v^2(a)} && \text{(mise au même dénominateur)}\end{aligned}$$

□

Théorème (Dérivation d'une fonction réciproque)**.** *Soit $f : I\to\mathbb{R}$ une fonction continue strictement monotone sur un intervalle I (et donc bijective entre I et l'intervalle $f(I)$).*
Si *f est dérivable en $a\in I$ et $f'(a)\neq 0$,*
alors *f^{-1} est dérivable au point $b:=f(a)$, et $\left(f^{-1}\right)'(b)=\dfrac{1}{f'\left(f^{-1}(b)\right)}$.*

Démonstration. On a par hypothèse :

$$\frac{f(t)-f(a)}{t-a}\underset{t\to a}{\longrightarrow}f'(a)\neq 0$$

donc par passage à l'inverse :

$$\frac{t-a}{f(t)-f(a)}\underset{t\to a}{\longrightarrow}\frac{1}{f'(a)}$$

On réalise alors le changement de variable $y=f(t)$ (c'est-à-dire $t=f^{-1}(y)$), en remarquant que :
- $b=f(a)$ signifie que $a=f^{-1}(b)$;
- par continuité de f^{-1} (voir théorème de la bijection monotone), on a bien $f^{-1}(y)\underset{y\to b}{\longrightarrow}f^{-1}(b)=a$.

ce qui donne :

$$\frac{f^{-1}(y)-f^{-1}(b)}{y-b}\underset{y\to b}{\longrightarrow}\frac{1}{f'\left(f^{-1}(b)\right)}$$

D'après la première définition de la dérivabilité, cette dernière phrase signifie précisément que f^{-1} est dérivable en b, et que $\left(f^{-1}\right)'(b)=\dfrac{1}{f'\left(f^{-1}(b)\right)}$. □

Définition (Dérivabilité à droite/à gauche)**.**
On dit que f est dérivable à droite (respectivement à gauche) en a lorsque le taux d'accroissement $\frac{f(t)-f(a)}{t-a}$ admet une limite finie lorsque $t \to a^+$ (respectivement lorsque $t \to a^-$).
Lorsque cette limite existe, on la note $f'_d(a)$ (respectivement $f'_g(a)$).

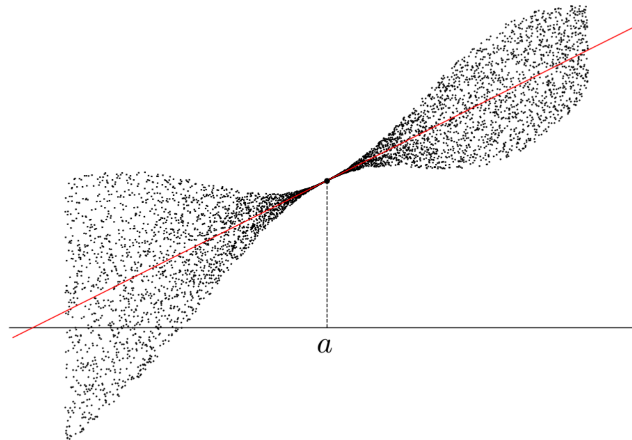
Proposition. *Si f est définie sur un voisinage $[a-\eta, a+\eta]$ autour de a, on a l'équivalence :*

$$(f \text{ est dérivable en } a) \iff \left(\begin{array}{c} f \text{ est dérivable à gauche } \textbf{et} \text{ à droite en } a \\ \text{et } f'_d(a) = f'_g(a) \end{array} \right)$$

Démonstration. En fait, rien à prouver ici. Cela tombe sous le sens par définition de la limite (limite en un point, limite à droite, limite à gauche). □

Remarque. *Tous les théorèmes et toutes les propositions donnés jusqu'ici sont également valables lorsqu'il s'agit non pas de dérivabilité, mais uniquement de dérivabilité à droite (ou respectivement de dérivabilité à gauche). Dans les démonstrations, il suffit de remplacer les limites $t \to a$ et $x \to 0$ par des limites $t \to a^+$ et $x \to 0^+$.*

Remarque. *Malgré les représentations traditionnelles, qui figurent toujours des fonctions très régulières sur tout un intervalle, il faut avoir à l'esprit qu'une fonction peut n'être dérivable qu'en un seul point a.*

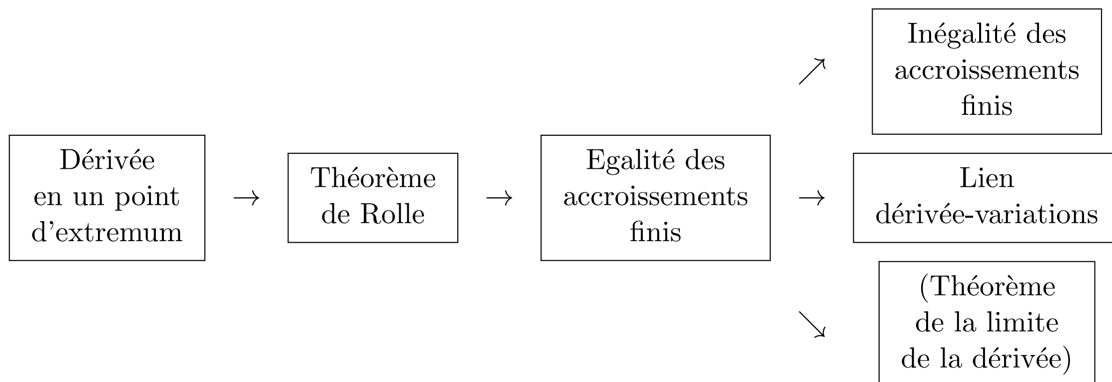


2 Dérivée sur un intervalle

Dans cette partie, I désigne un intervalle de $\mathbb{R}$.

Définition. *On dit que $f : I \to \mathbb{R}$ est dérivable sur I si elle l'est* ***en tout point de*** *I.*
Dans ce cas, la fonction qui à chaque $t \in I$ associe le nombre dérivé de f en t est notée f', et s'appelle la fonction dérivée de f.

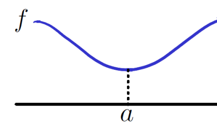
NB : à partir de maintenant et jusqu'à la fin de cette section, le plan du cours est globalement le suivant :



Le théorème de la limite de la dérivée a été écrit entre parenthèses, car il sera vu dans la chapitre sur les prolongements de fonctions.

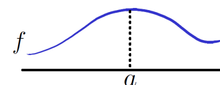
Définition (point d'extremum local)**.**
On dit qu'une fonction f admet un ***minimum local*** *en un point a de son ensemble de définition $\mathcal{D}_f$ s'il existe un voisinage $V_a = [a-\eta, a+\eta]$ de a tel que :*

$$\forall t \in V_a \cap \mathcal{D}_f \ , \ f(t) \geqslant f(a).$$

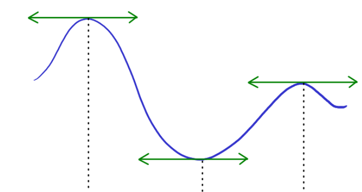


Symétriquement, on dit que a est un point de ***maximum local*** *de f s'il existe un voisinage V_a de a tel que :*

$$\forall t \in V_a \cap \mathcal{D}_f \ , \ f(t) \leqslant f(a).$$



Proposition (dérivée en un point d'extremum local)**.**
Soit f une fonction dérivable sur I, et a un point de ***l'intérieur de*** *I (c-à-d pas un bord de I).*
Si *a est un point d'extremum local (minimum ou maximum),* ***alors*** *$f'(a) = 0$.*



Démonstration. Supposons que $a \in I$ est un point d'extremum local placé à l'intérieur de I. Quitte à considérer la fonction $-f$ plutôt que f (ce qui ne change rien au problème), on peut supposer que a est un point de minimum local. Montrons que $f'(a) = 0$.
On sait que l'accroissement $P_a(t) := \frac{f(t)-f(a)}{t-a}$ tend vers $f'(a)$ lorsque $t \to a$. Or :

- sur un voisinage à droite de a, $P_a(t)$ est positif car :

$$\begin{cases} f(t) - f(a) \geqslant 0 & \text{(car } a \text{ est un point de minimum local)} \\ t - a > 0 & \text{(car } t > a) \end{cases}$$

Donc par passage à la limite $t \to a^+$ dans l'inégalité $P_a(t) \geqslant 0$, on a $\boxed{f'(a) \geqslant 0}$.

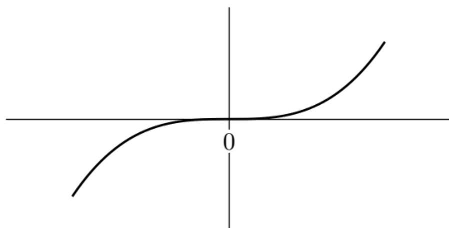
- sur un voisinage à gauche de a, $P_a(t)$ est négatif car :

$$\begin{cases} f(t) - f(a) \geqslant 0 & \text{(car } a \text{ est un point de minimum local)} \\ t - a < 0 & \text{(car } t < a\text{)} \end{cases}$$

Donc par passage à la limite $t \to a^-$ dans l'inégalité $P_a(t) \leqslant 0$, on a $\boxed{f'(a) \leqslant 0}$.

Finalement, on a bien : $f'(a) = 0$. □

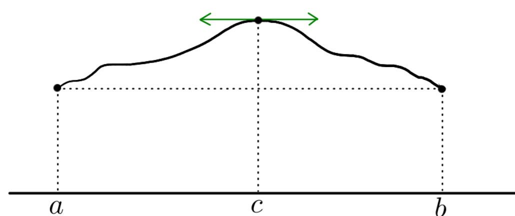
Vocabulaire. *Un point a tel que $f'(a) = 0$ est appelé un **point critique** de f. Ce n'est pas forcément un extremum local de f. Par exemple, la fonction $t \mapsto t^3$ a une dérivée nulle en 0, mais 0 n'est ni un minimum local, ni un maximum local.*



Théorème (Théorème de Rolle)**.**

Si $f : [a, b] \to \mathbb{R}$ (où $a < b$) est une fonction $\left|\begin{array}{l} \textit{continue sur } [a,b] \\ \textit{dérivable sur }]a,b[\\ \textit{telle que } f(a) = f(b) \end{array}\right.$

alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.



Démonstration. Considérons une telle fonction f.
NB : l'idée de la démonstration est de montrer qu'elle admet un point d'extremum local dans $]a, b[$, et d'appliquer la proposition précédente.

Comme f est continue sur le segment $[a, b]$, on sait d'après le théorème des bornes atteintes que f est bornée et atteint ses bornes. Autrement dit :

- il existe $u \in [a, b]$ tel que f réalise en u un minimum global (mettons $m = f(u)$) ;
- il existe $v \in [a, b]$ tel que f réalise en v un maximum global (mettons $M = f(v)$).

Premier cas : si l'une des deux valeurs u ou v n'est pas au bord de $[a, b]$ (mettons $u \in]a, b[$), alors u est un point d'extremum local (car global) à l'intérieur de $[a, b]$, et donc d'après la proposition précédente, $\boxed{f'(u) = 0}$.

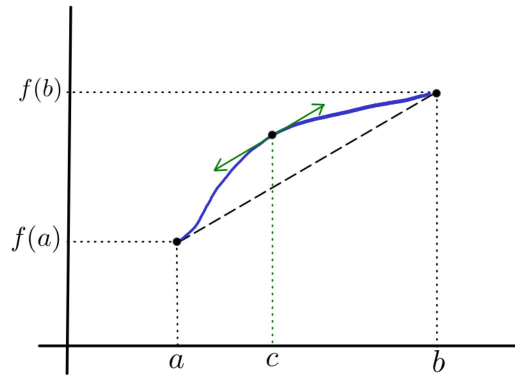
Deuxième cas : si les deux valeurs u et v sont sur le bords de $[a, b]$ (c'est-à-dire valent a ou b), alors on a :

$$f(u) = f(v) = f(a) = f(b) \qquad \text{et donc} \qquad m = M.$$

La valeur maximale de f est donc égale à sa valeur minimale, ce qui signifie que f est constante sur $[a, b]$. Dans ce cas, f' est la fonction nulle, et est donc nulle en n'importe quel point $c \in]a, b[$. □

Théorème (Egalité des accroissements finis)**.**
Si $f : [a, b] \to \mathbb{R}$ *(où* $a < b$*) est une fonction* $\left| \begin{array}{l} \textit{continue sur } [a,b] \\ \textit{dérivable sur }]a,b[\end{array} \right.$

alors il existe $c \in]a, b[$ *tel que* $f'(c) = \frac{f(b)-f(a)}{b-a}$ *(c'est-à-dire l'accroissement de* f *entre* a *et* b*).*



Démonstration. Considérons une telle fonction f.
NB : l'idée de la démonstration est d'appliquer le thm de Rolle à une fonction g *intelligemment modifiée de façon à avoir* $g(a) = g(b)$.

Notons $p = \frac{f(b)-f(a)}{b-a}$. On pose la fonction g définie sur $[a, b]$ par :

$$g(t) = f(t) - \big[f(a) + p(t - a)\big]$$

NB : g *est la différence entre* f *et la fonction affine qui relie les points* $A(a, f(a))$ *et* $B(b, f(b))$.

Cette fonction g vérifie $g(a) = g(b) = 0$ car :

$$g(a) = f(a) - \big[f(a) + 0\big] = 0$$

et

$$g(b) = f(b) - \big[f(a) + p(b - a)\big] = f(b) - \big[f(a) + f(b) - f(a)\big] = 0$$

g est donc une fonction $\left| \begin{array}{l} \text{continue sur } [a,b] \text{ (comme somme de fonctions continues)} \\ \text{dérivable sur }]a,b[\text{ (comme somme de fonctions dérivables)} \\ \text{telle que } g(a) = g(b) \end{array} \right.$.

D'après le théorème de Rolle, il existe $c \in]a, b[$ tel que $g'(c) = 0$. Or, en revenant à la définition de g et en dérivant, on a $g'(c) = f'(c) - p$. Donc $\boxed{f'(c) = p}$. □

Remarques.

1. *Ce théorème est une généralisation du théorème de Rolle, car dans le cas particulier où* $f(a) = f(b)$*, l'accroissement* $\frac{f(b)-f(a)}{b-a}$ *vaut* 0.
2. *Il signifie géométriquement que le graphe de* f *admet entre* a *et* b *une tangente parallèle à la droite reliant les points du plan* $A(a, f(a))$ *et* $B(b, f(b))$.
3. *Il y a un lien double entre les accroissements d'une fonction dérivable* f *d'une part, et ses nombres dérivés* $f'(t)$ *d'autre part :*
   - *tout nombre dérivé* $f'(t)$ *est toujours la limite d'accroissements, par définition ;*
   - *tout accroissement* $\frac{f(b)-f(a)}{b-a}$ *est le nombre dérivé* $f'(c)$ *pour un point* $c \in]a, b[$*, d'après l'égalité des accroissements finis.*

Théorème. *Soit $f : I \to \mathbb{R}$ une fonction dérivable (I est un intervalle). On a les équivalences suivantes :*

1. $\big(f \text{ croissante}\big) \iff \big(\forall x \in I\,,\ f'(x) \geqslant 0\big)$
2. $\big(f \text{ décroissante}\big) \iff \big(\forall x \in I\,,\ f'(x) \leqslant 0\big)$
3. $\big(f \text{ constante}\big) \iff \big(\forall x \in I\,,\ f'(x) = 0\big)$

Démonstration.

1. - Montrons ($\Longrightarrow$).
     Supposons que f est croissante. Alors tous les accroissements $\frac{f(x)-f(y)}{x-y}$ sont positifs. En effet, si $x > y$, alors $\frac{f(x)-f(y)}{x-y}$ est un quotient de nombres positifs, et si $x < y$, c'est un quotient de nombres négatifs (donc positif). Donc si on fixe $x \in I$ et que l'on fait un passage à la limite $y \to x$ dans l'inégalité :
     $$\frac{f(x)-f(y)}{x-y} \geqslant 0$$
     On obtient $\boxed{f'(x) \geqslant 0}$.
   - Montrons ($\Longleftarrow$).
     Supposons que f' est positive sur I. Montrons que f est croissante. Pour cela, fixons deux points $a < b$ de I. L'objectif est donc de montrer que $f(a) \leqslant f(b)$.
     On remarque que c'est équivalent à démontrer que l'accroissement $\frac{f(b)-f(a)}{b-a}$ est positif, car comme son dénominateur est strictement positif, cet accroissement a le même signe que son numérateur. Or, par l'égalité des accroissements finis, il existe $c \in]a, b[$ tel que :
     $$\frac{f(b)-f(a)}{b-a} = f'(c)$$
     et par hypothèse, $f'(c) \geqslant 0$. Donc $\frac{f(b)-f(a)}{b-a} \geqslant 0$, ce qui équivaut à $\boxed{f(a) \leqslant f(b)}$.
2. Plutôt que de recommencer et adapter la démonstration précédente, on peut utiliser le 1 sur la fonction $-f$, en écrivant les équivalences suivantes :
   $$\begin{aligned}\big(f \text{ décroissante}\big) &\iff \big(-f \text{ croissante}\big)\\ &\iff \big((-f)' \geqslant 0\big) \qquad \text{(d'après 1)}\\ &\iff \big(f' \leqslant 0\big)\end{aligned}$$
3. Il s'agit d'une conséquence de 1 et 2 car une fonction constante est précisément une fonction qui est à la fois croissante et décroissante.

□

Théorème. *Soit $f : I \to \mathbb{R}$ une fonction dérivable (I est un intervalle).*
Si f' est positive (resp. négative) sur I, et ne s'annule qu'en un nombre fini de points de I, alors f est strictement croissante (resp. strictement décroissante) sur I.

Démonstration. Quitte à considérer la fonction $-f$ plutôt que f, on peut ne considérer que le cas où f' est positive. Considérons donc une fonction f qui vérifie les hypothèses du théorème.
Montrons que f est strictement croissante sur I. Pour cela, fixons deux points $a < b$ de I. L'objectif est donc de montrer que $f(a) < f(b)$.

NB : l'idée de la démonstration est d'adapter la démonstration précédente (la partie ($\Longleftarrow$)) en s'arrangeant pour appliquer l'égalité des accroissements finis à un endroit où f' est strictement positive. Il s'agit donc d'éviter les quelques points problématiques où f' s'annule.

On note $x_0 < x_1 < \ldots < x_n < x_{n+1}$ les points de I tels que :

$$\begin{cases} x_0 = a \\ x_{n+1} = b \\ x_1, \ldots, x_n \text{ sont les valeurs d'annulation de } f' \text{ sur }]a, b[. \end{cases}$$

Tout d'abord, comme f' est positive, f est croissante, donc on a :

$$\underbrace{f(x_0)}_{=f(a)} \leqslant f(x_1) \leqslant \ldots \leqslant f(x_n) \leqslant \underbrace{f(x_{n+1})}_{=f(b)} \qquad (*)$$

D'autre part, on peut appliquer l'égalité des accroissements finis entre n'importe quel couple de points du type (x_i, x_{i+1}), où $i \in [\![0, n]\!]$: il existe $c_i \in]x_i, x_{i+1}[$ tel que

$$\frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} = f'(c_i)$$

Comme c_i ne vaut aucun des x_k (et donc aucune des valeurs d'annulation de f'), on a $f'(c_i) > 0$. Donc $\frac{f(x_{i+1})-f(x_i)}{x_{i+1}-x_i} > 0$, ce qui signifie que $f(x_i) < f(x_{i+1})$.
Les inégalités dans $(*)$ sont donc strictes, et donc $\boxed{f(a) < f(b)}$. □

Définition (Fonction lipschitzienne)**.**
Soit $M \geqslant 0$. On dit qu'une fonction f définie sur un intervalle I est M-lipschitzienne si pour tous $x, y \in I$:

$$|f(x) - f(y)| \ \leqslant \ M\,|x - y|$$

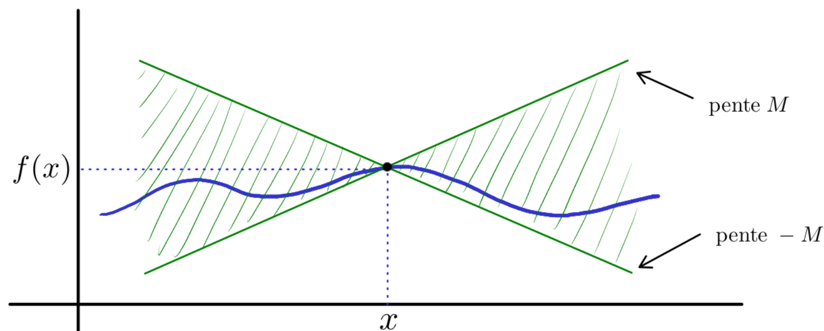
Remarques.

1. *Cela signifie que pour tout $x, y \in I$, la distance entre $f(x)$ et $f(y)$ (qui se lit sur l'axe des ordonnées) peut être majorée proportionnellement à la distance $|x - y|$ (qui se lit sur l'axe des abscisses).*
2. *C'est équivalent à dire que pour tous $x \neq y$ de I :*

$$\left|\frac{f(x) - f(y)}{x - y}\right| \leqslant M$$

Autrement dit, une fonction est M-lipschitzienne ssi ses accroissements sont bornés par M.

3. *Géométriquement, cela signifie que quel que soit $x \in I$, le graphe de f est inclus dans le cône suivant :*



Théorème (Inégalité des accroissements finis)**.** *Soit f une fonction dérivable sur un intervalle I.*
***Si** sa dérivée f' est bornée par un réel $M \geqslant 0$ (c'est-à-dire qu'on a $|f'(t)| \leqslant M$ pour tout $t \in I$),*
***alors** f est M-lipschitzienne.*

Démonstration. Considérons une fonction f dérivable sur I, et telle que $|f'| \leqslant M$.
Fixons $x, y \in I$, tels que $x \neq y$. Montrons que $\left|\frac{f(x)-f(y)}{x-y}\right| \leqslant M$.

Par l'égalité des accroissements finis, il existe $c \in]x, y[$ tel que $\frac{f(x)-f(y)}{x-y} = f'(c)$.
Or, par hypothèse, $|f'(c)| \leqslant M$.
Donc $\left|\frac{f(x)-f(y)}{x-y}\right| \leqslant M$, ce qui est bien équivalent à $|f(x) - f(y)| \leqslant M\,|x - y|$. □

Une application classique de l'inégalité des accroissements finis :
Pour montrer la convergence d'une suite u vérifiant $u_{n+1} = f(u_n)$, pour une certaine fonction f, une méthode efficace est de mettre en évidence un intervalle I tel que :

$$\begin{cases} I \text{ est stable par } f\text{, et contient un terme de } u \text{ (mettons } u_0\text{, pour simplifier)} ; \\ f \text{ a un unique point fixe } \ell \text{ sur } I ; \\ f \text{ est } k\text{-lipschitzienne pour un } k \in [0,1[\text{ (il est important que } k \text{ soit strictement inférieur à 1).} \end{cases}$$

Cette situation assure que $u_n \underset{n\to+\infty}{\longrightarrow} \ell$. En effet :

- ▶ Le fait que $u_0 \in I$ et que I est stable par f assure que tous les termes de u appartiennent à I, par récurrence immédiate.
- ▶ On peut alors écrire la chose suivante, pour tout $n \in \mathbb{N}$:

$$|u_{n+1} - \ell| = |f(u_n) - f(\ell)| \leqslant k|u_n - \ell|$$

  La première égalité est due au fait que $u_{n+1} = f(u_n)$, et $f(\ell) = \ell$ (puisque ℓ est un point fixe de f). L'inégalité qui suit est due au fait que f est k-lipschitzienne sur I. En effet, on a simplement appliqué l'inégalité $|f(x) - f(y)| \leqslant k|x - y|$ à $x = u_n$ et $y = \ell$ (qui appartiennent bien à I).
- ▶ L'inégalité $|u_{n+1} - \ell| \leqslant k|u_n - \ell|$, vraie pour tout n, permet d'avoir par récurrence que

$$\forall n \in \mathbb{N}\,,\ |u_n - \ell| \leqslant k^n |u_0 - \ell|.$$

- ▶ Comme $k \in [0,1[$, on a $k^n \underset{n\to+\infty}{\longrightarrow} 0$.
  Donc par comparaison, $|u_n - \ell| \underset{n\to+\infty}{\longrightarrow} 0$, c'est-à-dire $u_n \underset{n\to+\infty}{\longrightarrow} \ell$.

3 Fonctions de classe $\mathcal{C}^k$

On fixe un intervalle I de $\mathbb{R}$.

Notations. *Soit $k \in \mathbb{N}^*$.*

- *L'ensemble des fonctions de I dans $\mathbb{R}$ qui sont k fois dérivables est noté $\mathcal{D}^k(I,\mathbb{R})$.*
- *Si $f \in \mathcal{D}^k(I,\mathbb{R})$, alors pour tout $p \in [\![1,k]\!]$, la dérivée p-ième de f est notée $f^{(p)}$. Par convention, $f^{(0)} = f$.*
- *On dit que $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^k$ sur I (ou que f appartient à l'ensemble $\mathcal{C}^k(I,\mathbb{R})$) si f est k fois dérivable **et** $f^{(k)}$ est continue sur I.*
- *On dit que f est de classe $\mathcal{C}^\infty$ sur I (ou que $f \in \mathcal{C}^\infty(I,\mathbb{R})$) si f est infiniment dérivable, c'est-à-dire si $f \in \mathcal{C}^k(I,\mathbb{R})$ pour tout $k \in \mathbb{N}^*$.*

Remarque. *On a les inclusions suivantes :*

$$\mathcal{C}^\infty(I,\mathbb{R}) \subset \ldots \subset \mathcal{C}^2(I,\mathbb{R}) \subset \mathcal{D}^2(I,\mathbb{R}) \subset \mathcal{C}^1(I,\mathbb{R}) \subset \mathcal{D}^1(I,\mathbb{R}) \subset \mathcal{C}^0(I,\mathbb{R})$$

Propriétés. *Soit $k \in \mathbb{N}$ et u et v deux fonctions de classe $\mathcal{C}^k$ sur I.*

1. ***Somme :*** *$(u+v) \in \mathcal{C}^k(I,\mathbb{R})$ et $\forall p \in [\![0,k]\!]$, $(u+v)^{(p)} = u^{(p)} + v^{(p)}$.*
2. ***Produit par une constante :*** *$\forall \lambda \in \mathbb{R}$, $\lambda u \in \mathcal{C}^k(I,\mathbb{R})$ et $\forall p \in [\![0,k]\!]$, $(\lambda u)^{(p)} = \lambda u^{(p)}$.*
3. ***Produit :*** *$uv \in \mathcal{C}^k(I,\mathbb{R})$ (et voir plus loin pour $(uv)^{(p)}$).*
4. ***Composée :*** *Si f est de classe $\mathcal{C}^k$ sur l'intervalle $u(I)$, alors $f \circ u \in \mathcal{C}^k(I,\mathbb{R})$ (et pas de jolie formule pour $(f \circ u)^{(p)}$).*
5. ***Réciproque :*** *Si $u \in \mathcal{C}^k(I,\mathbb{R})$ réalise une bijection de I dans $u(I)$ et si u' ne s'annule jamais sur I, alors $u^{-1} \in \mathcal{C}^k(u(I),\mathbb{R})$.*

Démonstration. Des récurrences !

1. Récurrence finie sur $p \in [\![0,k]\!]$ en posant :

$$\mathcal{H}(p) \ :\ \text{" } (u+v) \in \mathcal{C}^p(I,\mathbb{R}) \text{ et } (u+v)^{(p)} = u^{(p)} + v^{(p)} \text{ "}$$

— Initialisation : $\mathcal{H}(0)$ signifie que $(u+v)$ est continue, et que $u+v=u+v$, ce qui est vrai.
— Hérédité : fixons $p \in [\![0, k-1]\!]$. Supposons $\mathcal{H}(p)$. On sait donc que :

$$(u+v)^{(p)} = u^{(p)} + v^{(p)} \qquad (*)$$

Or, u et v sont de classe $\mathcal{C}^k$, et donc à plus forte raison de classe $\mathcal{C}^{p+1}$. Donc $u^{(p)}$ et $v^{(p)}$ sont de classe $\mathcal{C}^1$. Par somme de fonctions dérivables, la fonction $(u+v)^{(p)}$ est donc dérivable, et en dérivant $(*)$, on obtient :

$$(u+v)^{(p+1)} = u^{(p+1)} + v^{(p+1)}$$

De plus, $u^{(p+1)}$ et $v^{(p+1)}$ sont continues, donc par somme de fonctions continues, $(u+v)^{(p+1)}$ est continue. On a donc bien prouvé que $(u+v)$ est de classe $\mathcal{C}^{p+1}$ et que $(u+v)^{(p+1)} = u^{(p+1)} + v^{(p+1)}$, c'est-à-dire $\mathcal{H}(p+1)$.
— Conclusion : pour tout $p \in [\![0, k]\!]$, $\mathcal{H}(p)$ est vraie.

2. Idem, en fixant $\lambda \in \mathbb{R}$ et en posant :

$$\mathcal{H}(p) \;:\; \text{" } \lambda u \in \mathcal{C}^p(I, \mathbb{R}) \text{ et } (\lambda u)^{(p)} = \lambda u^{(p)} \text{ "}$$

3. Si on ne s'intéresse qu'au caractère $\mathcal{C}^k$ de uv (et pas à la formule qui donne $(uv)^{(k)}$), on peut procéder par récurrence finie sur p, en disant dans l'hérédité que uv est de classe $\mathcal{C}^{p+1}$ si et seulement si sa dérivée $(uv)'$ est de classe $\mathcal{C}^p$. Or :

$$(uv)' = u'v + uv'$$

On voit alors qu'on a besoin d'appliquer l'hypothèse de récurrence non pas simplement à u et v mais aussi à u' et v'. La récurrence ne marche donc pas si l'on pose simplement :

$$\mathcal{H}(p) \;:\; \text{" } uv \in \mathcal{C}^p(I, \mathbb{R}) \text{ "}$$

pour deux fonctions u et v fixées. Il convient plutôt de poser :

$$\mathcal{H}(p) \;:\; \text{" } \textbf{Pour toutes fonctions } \mathbf{u}, \mathbf{v} \in \mathcal{C}^{\mathbf{p}}(\mathbf{I}, \mathbb{R}), \text{ on a } uv \in \mathcal{C}^p(I, \mathbb{R}) \text{ "}$$

A méditer !

4. Même difficulté que précédemment. Dans l'hérédité, on peut dire que $f \circ u$ est de classe $\mathcal{C}^{p+1}$ si et seulement si $(f \circ u)'$ est de classe $\mathcal{C}^p$. Or :

$$(f \circ u)' = (f' \circ u) \times u'$$

On a un produit (ça ok grâce à 3) mais en revanche la composition ne concerne pas f mais f'. On peut donc s'en tirer en fixant u de classe $\mathcal{C}^k$ et en posant pour tout $p \in [\![0, k]\!]$:

$$\mathcal{H}(p) \;:\; \text{" } \textbf{Pour toute fonction } \mathbf{f} \in \mathcal{C}^{\mathbf{p}}(\mathbf{u}(\mathbf{I}), \mathbb{R}), \text{ on a } f \circ u \in \mathcal{C}^p(I, \mathbb{R}) \text{ "}$$

5. Ici, on peut faire une récurrence sur p pour une fonction u fixée, en posant simplement :

$$\mathcal{H}(p) \;:\; \text{" } u^{-1} \in \mathcal{C}^p(u(I), \mathbb{R}) \text{ "}$$

En effet, dans l'hérédité, on peut dire que u^{-1} est de classe $\mathcal{C}^{p+1}$ ssi sa dérivée $(u^{-1})'$ est de classe $\mathcal{C}^p$. Or :

$$(u^{-1})' = \frac{1}{u' \circ u^{-1}}$$

On reconnaît alors une composition de fonctions de classe $\mathcal{C}^p$, donc c'est bon. Rigoureusement parlant, comme cette composition fait intervenir la fonction inverse, il convient de montrer préalablement que $f \;:\; t \mapsto \frac{1}{t}$ est de classe $\mathcal{C}^\infty$. Cela peut être fait par récurrence en posant pour tout $n \in \mathbb{N}$:

$$\mathcal{P}(n) \;:\; \text{" } f \text{ est } \mathcal{C}^n \text{ sur } \mathbb{R}_-^* \text{ et } \mathbb{R}_+^*, \text{ et } \forall t \in \mathbb{R}^*,\; f^{(n)}(t) = \frac{(-1)^n n!}{t^{n+1}} \text{ "}$$

□

Remarque. *Comme ces propriétés sont vraies pour tout $k \in \mathbb{N}$, Une somme/produit/composée/réciproque de fonctions de classe $\mathcal{C}^\infty$ est de classe $\mathcal{C}^\infty$.*

Théorème (Formule de Leibniz)**.** *Soit $n \in \mathbb{N}$ et u, v des fonctions de I dans $\mathbb{R}$. Si u et v sont n fois dérivables, alors :*

$$(uv)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} u^{(k)} v^{(n-k)}$$

Démonstration. On raisonne par récurrence, en posant pour tout $n \in \mathbb{N}$:

$$\mathcal{H}(n) : \text{" si } u \text{ et } v \text{ sont } n \text{ fois dérivables, alors } (uv)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} u^{(k)} v^{(n-k)} \text{ "}$$

— Initialisation. $\mathcal{H}(0)$ signifie que $uv = \binom{0}{0} u^{(0)} v^{(0)}$, ce qui est vrai.

— Hérédité. Fixons $n \in \mathbb{N}$. Supposons $\mathcal{H}(n)$. En vue de démontrer $\mathcal{H}(n+1)$, supposons que u et v sont $(n+1)$ fois dérivables sur I. L'objectif est alors de démontrer que

$$(uv)^{(n+1)} = \sum_{k=0}^{n+1} \binom{n+1}{k} u^{(k)} v^{(n+1-k)}$$

Pour cela, on écrit :

$$\begin{aligned}
(uv)^{(n+1)} &= \left((uv)^{(n)}\right)' \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} u^{(k)} v^{(n-k)}\right)' && \text{(d'après } \mathcal{H}(n)\text{)} \\
&= \sum_{k=0}^{n} \binom{n}{k} \left(u^{(k)} v^{(n-k)}\right)' && \text{(par linéarité de la dérivation)} \\
&= \sum_{k=0}^{n} \binom{n}{k} \left(u^{(k+1)} v^{(n-k)} + u^{(k)} v^{(n-k+1)}\right) && \text{(dérivation du produit)} \\
&= \sum_{k=0}^{n} \binom{n}{k} u^{(k+1)} v^{(n-k)} + \sum_{k=0}^{n} \binom{n}{k} u^{(k)} v^{(n-k+1)} \\
&= \sum_{p=1}^{n+1} \binom{n}{p-1} u^{(p)} v^{(n+1-p)} + \sum_{k=0}^{n} \binom{n}{k} u^{(k)} v^{(n+1-k)} && (p = k+1 \text{ dans la 1ère somme})
\end{aligned}$$

A ce stade, on peut remarquer que les deux sommes peuvent aller de 0 à $n+1$ car les termes manquants (les termes en $p = 0$ et en $k = n+1$) sont nuls. D'autre part, en renommant l'indice p en k, on voit que les deux sommes sont des combinaisons linéaires des $u^{(k)} v^{(n+1-k)}$. La réunion sous une seule somme donne donc :

$$(uv)^{(n+1)} = \sum_{k=0}^{n+1} \left[\binom{n}{k-1} + \binom{n}{k}\right] u^{(k)} v^{(n+1-k)}$$

Enfin, par la formule de Pascal, l'intérieur du crochet vaut $\binom{n+1}{k}$.
$\mathcal{H}(n+1)$ est bien démontrée.

— Conclusion : pour tout $n \in \mathbb{N}$, $\mathcal{H}(n)$ est vraie.

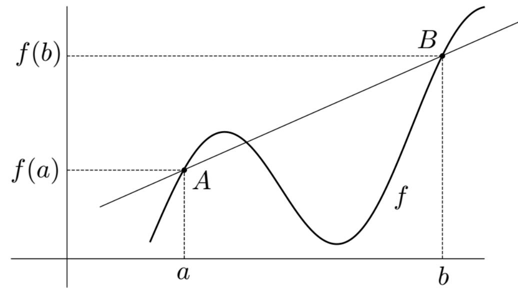
□

4 Fonctions convexes, fonctions concaves

On fixe un intervalle I de $\mathbb{R}$ et f une fonction de I dans $\mathbb{R}$.

Vocabulaire.

- *Une **droite sécante** (ou **sécante**) au graphe de f est une droite (AB) passant par deux points distincts du graphe $A(a, f(a))$ et $B(b, f(b))$, où $a, b \in I$, avec $a \neq b$.*



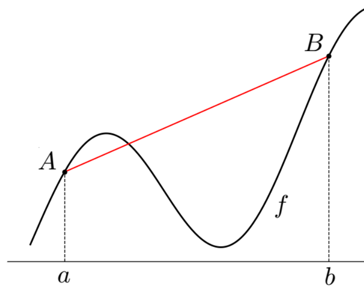
Cette droite a pour pente $\frac{f(b)-f(a)}{b-a}$ et passe par A, donc elle est portée par la fonction affine :

$$S_{a,b} \,:\, x \mapsto f(a) + \left(\frac{f(b)-f(a)}{b-a}\right)(x-a)$$

qui, en remarquant que la droite passe aussi par B, peut aussi s'écrire de la façon suivante :

$$S_{a,b} \,:\, x \mapsto f(b) + \left(\frac{f(b)-f(a)}{b-a}\right)(x-b)$$

- *Dans ce contexte, le segment $[AB]$ est appelé une **corde** du graphe de f.*



Définition (Première définition d'une fonction convexe)**.**
*$f \,:\, I \to \mathbb{R}$ est dite **convexe** si pour tous $a, b \in I$:*

$$\forall \lambda \in [0,1]\,, \quad f\big((1-\lambda)a + \lambda b\big) \leqslant (1-\lambda)f(a) + \lambda f(b)$$

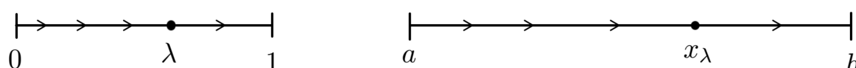
Remarque fondamentale

Fixons $a, b \in I$. Pour tout $\lambda \in [0,1]$, posons :

$$\begin{aligned} x_\lambda &= (1-\lambda)a + \lambda b \\ &= a + \lambda(b-a) \end{aligned}$$

Sous cette dernière écriture, on remarque que lorsque λ parcourt de gauche à droite l'intervalle $[0,1]$, le nombre x_λ parcourt l'intervalle $[a,b]$, avec notamment :

$$x_0 = a \qquad \text{et} \qquad x_1 = a + (b-a) = b$$



- Les $f\big((1-\lambda)a+\lambda b\big) = f(x_\lambda)$ où $\lambda \in [0,1]$ sont donc les valeurs que prend la fonction f sur l'intervalle $[a,b]$.
- De l'autre côté, lorsque λ parcourt $[0,1]$, la quantité $(1-\lambda)f(a)+\lambda f(b) = f(a)+\lambda(f(b)-f(a))$ prend toutes les valeurs intermédiaires entre $f(a)$ et $f(b)$. On peut l'exprimer en fonction de x_λ, car en inversant la formule donnant x_λ en fonction de λ, on obtient λ en fonction de x_λ :

$$\lambda = \frac{x_\lambda - a}{b-a}$$

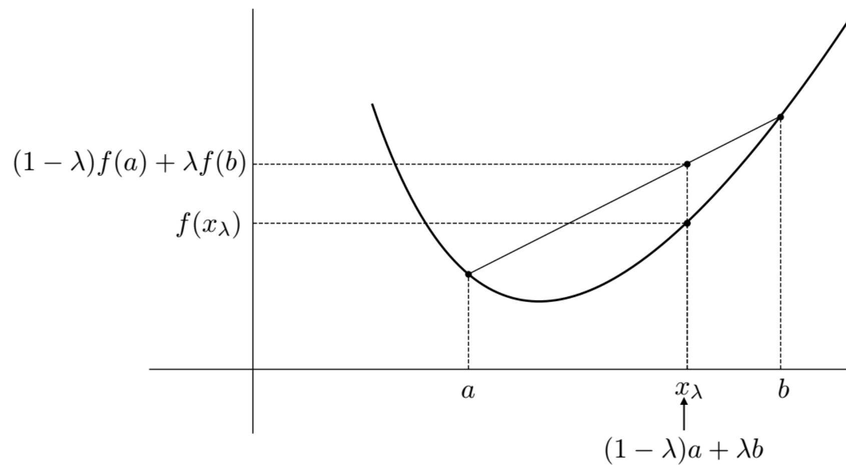
d'où

$$\begin{aligned}(1-\lambda)f(a)+\lambda f(b) &= f(a)+\lambda(f(b)-f(a))\\ &= f(a)+\frac{x_\lambda - a}{b-a}(f(b)-f(a))\\ &= f(a)+\left(\frac{f(b)-f(a)}{b-a}\right)(x_\lambda - a)\\ &= S_{a,b}(x_\lambda)\end{aligned}$$

Il s'agit de la sécante au graphe de f entre a et b !

D'où les équivalences suivantes :

$$\begin{aligned}&\forall\lambda\in[0,1]\,,\ f\big((1-\lambda)a+\lambda b\big) \leqslant (1-\lambda)f(a)+\lambda f(b)\\ \Longleftrightarrow\ &\forall\lambda\in[0,1]\,,\ f(x_\lambda)\leqslant S_{a,b}(x_\lambda)\\ \Longleftrightarrow\ &\forall x\in[a,b]\,,\ f(x)\leqslant S_{a,b}(x)\\ \Longleftrightarrow\ &\text{Le graphe de } f \text{ est sous sa corde entre } a \text{ et } b.\end{aligned}$$

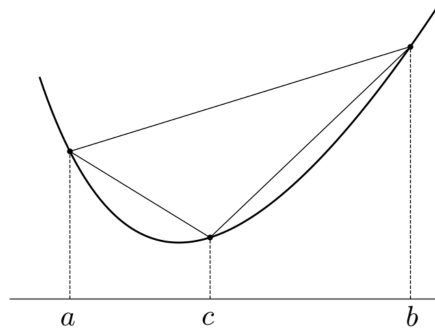


Définition (Seconde définition d'une fonction convexe, équivalente à la première)**.**
$f : I \to \mathbb{R}$ est dite convexe si son graphe est situé au-dessous de toutes ses cordes.

Lemme (Lemme des pentes, hors programme, mais très pratique dans les prochaines démonstrations.)**.**

Si f est convexe sur I, alors pour tout triplet de réels $a < c < b$ de I, on a :

$$\frac{f(c)-f(a)}{c-a} \leqslant \frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(c)}{b-c}$$



Démonstration. Fixons $a < c < b$ trois réels de I et procédons par équivalences :

$$\begin{aligned}\frac{f(c)-f(a)}{c-a} \leqslant \frac{f(b)-f(a)}{b-a} &\iff f(c) \leqslant f(a) + \left(\frac{f(b)-f(a)}{b-a}\right)(c-a)\\ &\iff f(c) \leqslant S_{a,b}(c)\,,\text{ ce qui est vrai par la def 2.}\end{aligned}$$

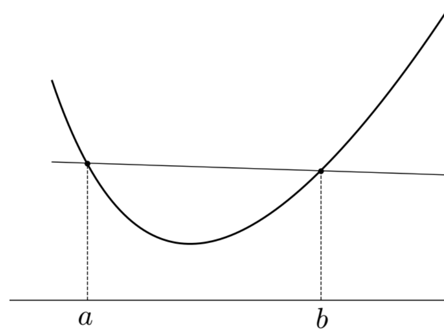
et de même :

$$\begin{aligned}\frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(c)}{b-c} &\iff f(c) \leqslant f(b) + \left(\frac{f(b)-f(a)}{b-a}\right)(c-b)\\ &\iff f(c) \leqslant S_{a,b}(c)\,,\text{ ce qui est vrai par la def 2.}\end{aligned}$$

Par remontée des équivalences, l'encadrement est bien vrai. □

Proposition. *Si f est convexe, alors pour tout $a, b \in I$:*

1. *$\forall x \in [a,b]$, $f(x) \leqslant S_{a,b}(x)$ (ça on le savait déjà) ;*
2. *$\forall x \in I \setminus [a,b]$, $f(x) \geqslant S_{a,b}(x)$ (f est au dessus de sa sécante à l'extérieur de $[a,b]$).*



Démonstration. Fixons $a < b$ deux réels de I. Seul le second point est à démontrer. Si $x > b$, l'application du lemme des pentes aux trois réels $a < b < x$ donne :

$$\frac{f(b)-f(a)}{b-a} \leqslant \frac{f(x)-f(a)}{x-a}$$

ce qui, en isolant $f(x)$, donne précisément $f(x) \geqslant S_{a,b}(x)$.
De même, si $x < a$, l'application du lemme des pentes aux trois réels $x < a < b$ donne :

$$\frac{f(b)-f(x)}{b-x} \leqslant \frac{f(b)-f(a)}{b-a}$$

ce qui, en isolant $f(x)$, donne précisément $f(x) \geqslant S_{a,b}(x)$. □

Théorème. *Supposons que f est dérivable sur I. Alors on a l'équivalence :*

$$f \text{ est convexe} \iff f' \text{ est croissante}$$

et si f est deux fois dérivable sur I, on a l'équivalence :

$$f \text{ est convexe} \iff f'' \geqslant 0$$

Démonstration. Il suffit de montrer la première équivalence, puisque la seconde en est une conséquence directe.

▶ Montrons l'implication : f est convexe $\implies$ f' est croissante.
Supposons f convexe. Fixons $a < b$ deux réels. L'objectif est de montrer que $f'(a) \leqslant f'(b)$. Par le lemme des pentes, on a pour tout $c \in]a, b[$:

$$\frac{f(c) - f(a)}{c - a} \leqslant \frac{f(b) - f(a)}{b - a} \leqslant \frac{f(b) - f(c)}{b - c}$$

Le passage à la limite $c \to a^+$ dans l'inégalité de gauche donne : $f'(a) \leqslant \frac{f(b)-f(a)}{b-a}$.
Le passage à la limite $c \to b^-$ dans l'inégalité de droite donne : $\frac{f(b)-f(a)}{b-a} \leqslant f'(b)$.
Finalement, par transitivité, $\boxed{f'(a) \leqslant f'(b)}$.

▶ Montrons l'implication : f' est croissante $\implies$ f est convexe.
Supposons que f' est croissante. Soit $a < b$ deux réels de I. On pose la fonction :

$$g : \left| \begin{array}{ccc} [a, b] & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) - S_{a,b}(x) \end{array} \right.$$

L'objectif est de démontrer que g est négative sur $[a, b]$. Par construction, $g(a) = g(b) = 0$. D'autre part, elle est dérivable, et pour tout $x \in [a, b]$:

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$$

D'après l'égalité des accroissements finis, il existe $c \in]a, b[$ tel que $\frac{f(b)-f(a)}{b-a} = f'(c)$. On a donc :

$$g'(x) = f'(x) - f'(c)$$

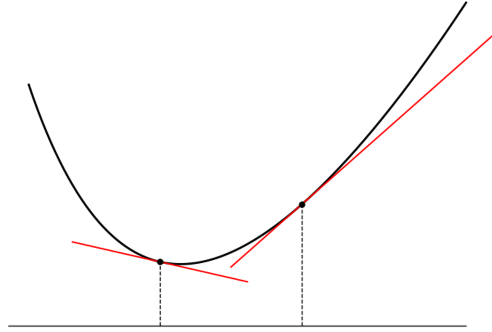
Par croissance de f' :

- g' est négative sur $[a, c]$, donc g est décroissante sur cet intervalle, et donc pour tout $x \in [a, c]$, $g(x) \leqslant g(a) = 0$;
- g' est positive sur $[c, b]$, donc g est croissante sur cet intervalle, et donc pour tout $x \in [c, b]$, $g(x) \leqslant g(b) = 0$.

Finalement, pour tout $x \in [a, b]$, $g(x) \leqslant 0$, c'est-à-dire $\boxed{f(x) \leqslant S_{a,b}(x)}$.
Ceci étant valable pour tous $a < b$ de I, f est donc convexe.

□

Théorème. *Si f est convexe et dérivable, alors son graphe est situé au-dessus de toutes ses tangentes.*



Démonstration. Supposons f dérivable et convexe, et fixons $a \in I$. On pose :

$$T_a : x \mapsto f(a) + f'(a)(x-a)$$

la fonction affine qui définit la tangente au graphe de f en a. L'objectif est alors de montrer que la fonction

$$h : \left| \begin{array}{ccc} I & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) - T_a(x) \end{array} \right.$$

est positive sur I. On peut remarquer tout de suite que $h(a) = 0$. De plus, h est dérivable, et pour tout $x \in I$:

$$h'(x) = f'(x) - f'(a)$$

Comme f est convexe, f' est croissante, donc $h'(x) \leqslant 0$ pour $x \leqslant a$ et $h'(x) \geqslant 0$ pour $x \geqslant a$.
h est donc décroissante, s'annule en a, puis croissante. Elle est donc positive sur I. □

Théorème/définition.
On dit que la fonction f est ***concave*** *sur I si son opposée $(-f)$ est convexe, ce qui est à équivalent à dire que :*

$$\forall a, b \in I\,,\ \forall \lambda \in [0,1]\,,\ f\big((1-\lambda)a + \lambda b\big) \geqslant (1-\lambda)f(a) + \lambda f(b)$$

ou encore à dire que le graphe de f est situé au-dessus de toutes ses cordes, ou encore, si f est dérivable, que f' est décroissante, ou encore, si f est deux fois dérivable, que f'' est négative.

En outre, le graphe d'une fonction concave est situé au-dessous de toutes ses tangentes.

